PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/47, 16/18, G01N 33/566, C12Q 1/68, C12N 15/11, 15/62, A01K 67/027, A61K 38/00

A2

(11) International Publication Number: WO 00/58473

(43) International Publication Date: 5 October 2000 (05.10.00)

(21) International Application Number: PCT/US00/08621

(22) International Filing Date: 31 March 2000 (31.03.00)

(30) Priority Data:

60/127,607	31 March 1999 (31.03.99)	US
60/127,636	2 April 1999 (02.04.99)	US
60/127,728	5 April 1999 (05.04.99)	US
09/540,763	30 March 2000 (30.03.00)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/127,607 (CIP)
Filed on	31 March 1999 (31.03.99)
US	60/127,636 (CIP)
Filed on	2 April 1999 (02.04.99)
US	60/127,728 (CIP)
Filed on	5 April 1999 (05.04.99)
US	09/540,763 (CIP)
Filed on	30 March 2000 (30.03.00)

(71) Applicant *(for all designated States except US)*: CURAGEN CORPORATION [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: SHIMKETS, Richard, A. [US/US]; 191 Leete Street, West Haven, CT 06516 (US). LEACH, Martin [GB/US]; 884 School Street, Webster, MA 01570 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: NUCLEIC ACIDS INCLUDING OPEN READING FRAMES ENCODING POLYPEPTIDES; "ORFX"

(57) Abstract

The present invention provides open reading frames ORFX, encoding isolated polypeptides, as well as polynucleotides encoding ORFX and antibodies that immunospecifically bind to ORFX or any derivative, variant, mutant, or fragment of the ORFX polypeptides, polynucleotides or antibodies. The invention additionally provides methods in which the ORFX polypeptide, polynucleotide and antibody are used in detection and treatment of a broad range of pathological states, as well as to other uses.

Column 5 of Table 1, entitled, "Cells or Tissues in Which Gene is Expressed", denotes tissues, represented by five digit numbers, in which RNA homologous to the ORF nucleic acid sequences is present. Tissues or cells corresponding to the numbers are provided in Table 2.

ORFX nucleic acids, and their encoded polypeptides, according to the invention are useful in a variety of applications and contexts. For example, various ORFX nucleic acids and polypeptides according to the invention are useful, *inter alia*, as novel members of the protein families indicated in Table 1, and/or according to the presence of domains and sequence relatedness to previously described proteins as summarized in Table 1.

ORFX nucleic acids and polypeptides according to the invention can also be used to identify cell types listed in Table 1 for an indicated ORFX according to the invention. Additional utilities for ORFX nucleic acids and polypeptides according to the invention are disclosed herein.

ORFX Nucleic Acids

The novel nucleic acids of the invention include those that encode an ORFX or ORFX-like protein, or biologically active portions thereof. The nucleic acids include nucleic acids encoding polypeptides that include the amino acid sequence of one or more of SEQ ID NO:2*n*, wherein n = 1 to 3161. The encoded polypeptides can thus include, *e.g.*, the amino acid sequences of SEQ ID NO: 2, 4, 6, 8, 10, . . ., 6310, 6312, 6314, 6316, 6318, 6320, and/or 6322.

In some embodiments, a nucleic acid encoding a polypeptide having the amino acid sequence of one or more of SEQ ID NO:2*n* (wherein n = 1 to 3161) includes the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein n = 1 to 3161), or a fragment thereof. Additionally, the invention includes mutant or variant nucleic acids of any of SEQ ID NO:2*n*-1 (wherein n = 1 to 3161), or a fragment thereof, any of whose bases may be changed from the disclosed sequence while still encoding a protein that maintains its ORFX -like activities and physiological functions. The invention further includes the complement of the nucleic acid sequence of any of SEQ ID NO:2*n*-1 (wherein n = 1 to 3161), including fragments, derivatives,

phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in ORFX is replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a ORFX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for ORFX biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

In one embodiment, a mutant ORFX protein can be assayed for (1) the ability to form protein:protein interactions with other ORFX proteins, other cell-surface proteins, or biologically active portions thereof, (2) complex formation between a mutant ORFX protein and a ORFX receptor; (3) the ability of a mutant ORFX protein to bind to an intracellular target protein or biologically active portion thereof; (*e.g.*, avidin proteins); (4) the ability to bind BRA protein; or (5) the ability to specifically bind an anti-ORFX protein antibody.

Antisense

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:2*n*-1 (wherein n = 1 to 3161), or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire ORFX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a ORFX protein of any of SEQ ID NO:2*n* (wherein n = 1 to 3161) or antisense nucleic acids complementary to a ORFX nucleic acid sequence of SEQ ID NO:2n-1 (wherein n = 1 to 3161) are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding ORFX. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid

the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

ORFX polypeptides

The novel protein of the invention includes the ORFX-like protein whose sequence is provided in any of SEQ ID NO:2*n* (wherein n = 1 to 3161). The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in FIG. 1 while still encoding a protein that maintains its ORFX-like activities and physiological functions, or a functional fragment thereof. For example, the invention includes the polypeptides encoded by the variant ORFX nucleic acids described above. In the mutant or variant protein, up to 20% or more of the residues may be so changed.

In general, an ORFX -like variant that preserves ORFX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above. Furthermore, without limiting the scope of the invention, positions of any of SEQ ID NO:2*n* (wherein n = 1 to 3161) may be substitute such that a mutant or variant protein may include one or more substitutions

The invention also includes isolated ORFX proteins, and biologically active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also provided are polypeptide fragments suitable for use as immunogens to raise anti-ORFX antibodies. In one embodiment, native ORFX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, ORFX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a ORFX

transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recrusive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify ORFX variants (Arkin and Yourvan (1992) PNAS 89:7811-7815; Delgrave *et al.* (1993) Protein Engineering 6:327-331).

Anti-ORFX Antibodies

The invention further encompasses antibodies and antibody fragments, such as F_{ab} or $(F_{ab})_2$, that bind immunospecifically to any of the proteins of the invention.

An isolated ORFX protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind ORFX using standard techniques for polyclonal and monoclonal antibody preparation. Full-length ORFX protein can be used. Alternatively, the invention provides antigenic peptide fragments of ORFX for use as immunogens. The antigenic peptide of ORFX comprises at least 4 amino acid residues of the amino acid sequence shown in any of SEQ ID NO:2*n* (wherein n = 1 to 3161). The antigenic peptide encompasses an epitope of ORFX such that an antibody raised against the peptide forms a specific immune complex with ORFX. The antigenic peptide may comprise at least 6 aa residues, at least 8 aa residues, at least 10 aa residues, at least 15 aa residues, at least 20 aa residues, or at least 30 aa residues. In one embodiment of the invention, the antigenic peptide comprises a polypeptide comprising at least 6 contiguous amino acids of any of SEQ ID NO:2*n* (wherein n = 1 to 3161).

In an embodiment of the invention, epitopes encompassed by the antigenic peptide are regions of ORFX that are located on the surface of the protein, *e.g.*, hydrophilic regions. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each incorporated herein by reference in their entirety.

As disclosed herein, an ORFX protein sequence of any of SEQ ID NO:2*n* (wherein n = 1 to 3161) , or derivatives, fragments, analogs or homologs thereof, may be utilized as immunogens in the generation of antibodies that immunospecifically-bind these protein components. The term "antibody" as used herein refers to immunoglobulin molecules and

phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

ORFX Recombinant Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding ORFX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements

the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding ORFX has been introduced) in a suitable medium such that ORFX protein is produced. In another embodiment, the method further comprises isolating ORFX from the medium or the host cell.

Transgenic animals

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which ORFX-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous ORFX sequences have been introduced into their genome or homologous recombinant animals in which endogenous ORFX sequences have been altered. Such animals are useful for studying the function and/or activity of ORFX and for identifying and/or evaluating modulators of ORFX activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous ORFX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing ORFX-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The human ORFX DNA sequence of SEQ ID NO:$2n$-1 (wherein n = 1 to 3161) can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of the human ORFX gene, such as a mouse ORFX gene, can be isolated based on hybridization to the human ORFX cDNA (described further above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of

example proliferative or differentiative disorders, or production of ORFX protein forms that have decreased or aberrant activity compared to ORFX wild type protein. In addition, the anti-ORFX antibodies of the invention can be used to detect and isolate ORFX proteins and modulate ORFX activity.

This invention further pertains to novel agents identified by the above described screening assays and uses thereof for treatments as described herein.

Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) that bind to ORFX proteins or have a stimulatory or inhibitory effect on, for example, ORFX expression or ORFX activity.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of a ORFX protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc Natl Acad Sci U.S.A.* 90:6909; Erb *et al.* (1994) *Proc Natl Acad Sci U.S.A.* 91:11422; Zuckermann *et al.* (1994) *J Med Chem* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2059; Carell *et al.* (1994) *Angew Chem Int Ed Engl* 33:2061; and Gallop *et al.* (1994) *J Med Chem* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), on chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner U.S. Pat. No. 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and

would abrogate the specific growth-inducing effects of the proteins. Such antibodies may be useful, for example, in the treatment of proliferative disorders including various tumors and benign hyperplasias.

Polynucleotides or oligonucleotides corresponding to any one portion of the ORFX nucleic acids of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161) may be used to detect DNA containing a corresponding ORF gene, or detect the expression of a corresponding ORFX gene, or ORFX-like gene. For example, an ORFX nucleic acid expressed in a particular cell or tissue, as noted in Table 2, can be used to identify the presence of that particular cell type.

An exemplary method for detecting the presence or absence of ORFX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes ORFX protein such that the presence of ORFX is detected in the biological sample. An agent for detecting ORFX mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to ORFX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length ORFX nucleic acid, such as the nucleic acid of SEQ ID NO:2*n*-1 (wherein *n* = 1 to 3161), or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to ORFX mRNA or genomic DNA, as described above. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting ORFX protein is an antibody capable of binding to ORFX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect ORFX mRNA, protein, or genomic DNA in a biological sample *in vitro* as

contracture, diabetic complications, or rheumatoid arthritis, etc.; and glia-associated disorders such as cerebral lesions, diabetic neuropathies, cerebral edema, senile dementia, Alzheimer's disease, etc. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant ORFX expression or activity in which a test sample is obtained from a subject and ORFX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant ORFX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant ORFX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder, such as a proliferative disorder, differentiative disorder, glia-associated disorders, etc. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant ORFX expression or activity in which a test sample is obtained and ORFX protein or nucleic acid is detected (*e.g.*, wherein the presence of ORFX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant ORFX expression or activity.)

The methods of the invention can also be used to detect genetic lesions in a ORFX gene, thereby determining if a subject with the lesioned gene is at risk for, or suffers from, a proliferative disorder, differentiative disorder, glia-associated disorder, etc. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding a ORFX-protein, or the mis-expression of the ORFX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of (1) a deletion of one or more nucleotides from a ORFX gene; (2) an addition of one or more nucleotides to a ORFX gene; (3) a substitution of one or more nucleotides of a ORFX gene, (4) a chromosomal

may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, EPIDERMAL WOUND HEALING, pp. 71-112 (Maibach and Rovee, eds.), Year Book Medical

<212> DNA
<213> Homo sapiens

<400> 2049
cgcgtcgctt acggtgcgct gaataccagc ctgctggcgc tggcggtcag cttcgcgtcg
60
ctgttcctcg ggatagtgtt cgggctgatg ccacgtctga tgtgcggggt gattgaactg
120
gccaacgctc cccgccaat cgccctgggc ctgttagtag tcgccattag cggccccttca
180
gcctacggtg ccgcctgtgc ggtgatgttg gtcagttggg ctccgctggc cgcccattgt
240
gcttcgttgt tggcggaagc ccgcacgcag ccctatatcc gcatgttgcc ggtattgggc
300
gtcggccgat ggcgcacgct gacccactac ctgctgccgg cgctctctgc tcccctgctg
360
cgccacgcca tgttgcgtct gccgggcatt gcgctggcgc tggcggcctt gggtttttt
420
ggtcttgggc cgcagccacc cagtgcagaa tgggggctgg tgctggcgga aggcatgcct
480
tatctcgaac gggcgccctg gggagtcctg gcaccg
516

<210> 2050
<211> 172
<212> PRT
<213> Homo sapiens

<400> 2050
Arg Val Ala Tyr Gly Ala Leu Asn Thr Ser Leu Leu Ala Leu Ala Val
1 5 10 15
Ser Phe Ala Ser Leu Phe Leu Gly Ile Val Phe Gly Leu Met Pro Arg
20 25 30
Leu Met Cys Gly Val Ile Glu Leu Ala Asn Ala Pro Pro Pro Ile Ala
35 40 45
Leu Gly Leu Leu Val Val Ala Ile Ser Gly Pro Ser Ala Tyr Gly Ala
50 55 60
Ala Cys Ala Val Met Leu Val Ser Trp Ala Pro Leu Ala Ala His Cys
65 70 75 80
Ala Ser Leu Leu Ala Glu Ala Arg Thr Gln Pro Tyr Ile Arg Met Leu
85 90 95
Pro Val Leu Gly Val Gly Arg Trp Arg Thr Leu Thr His Tyr Leu Leu
100 105 110
Pro Ala Leu Ser Ala Pro Leu Leu Arg His Ala Met Leu Arg Leu Pro
115 120 125
Gly Ile Ala Leu Ala Leu Ala Ala Leu Gly Phe Phe Gly Leu Gly Pro
130 135 140
Gln Pro Pro Ser Ala Glu Trp Gly Leu Val Leu Ala Glu Gly Met Pro
145 150 155 160
Tyr Leu Glu Arg Ala Pro Trp Gly Val Leu Ala Pro
165 170

<210> 2051
<211> 411

<212> DNA
<213> Homo sapiens

<400> 2051
gagcaaaact atcgttctac cggcaatatt ctgaaaagtg ccaaccaact tatttcgaat
60
aatagtgatc gtctcggtaa gaatttatgg accgacggtg aaatgggga gccagtaggt
120
atttatgcag catttaatga attagatgag gcaaaatttg tggcgtctca aatccaaaat
180
tgggtagatg atggtgggga attagatgat tgtgctgttt tatatcgtag taatagccaa
240
tctcgtgtta ttgaagaagc cttgattcgt tgccaaattc cttatcgaat ttatggcggg
300
atgcgattct tcgaacgcca agaaattaaa gatgcgttgg catatttacg tttaattaat
360
aatcgtcaag atgatgccgc atttgagcgt gtgattaata cgcctacgcg t
411

<210> 2052
<211> 137
<212> PRT
<213> Homo sapiens

<400> 2052
Glu Gln Asn Tyr Arg Ser Thr Gly Asn Ile Leu Lys Ser Ala Asn Gln
1 5 10 15
Leu Ile Ser Asn Asn Ser Asp Arg Leu Gly Lys Asn Leu Trp Thr Asp
20 25 30
Gly Glu Met Gly Glu Pro Val Gly Ile Tyr Ala Ala Phe Asn Glu Leu
35 40 45
Asp Glu Ala Lys Phe Val Ala Ser Gln Ile Gln Asn Trp Val Asp Asp
50 55 60
Gly Gly Glu Leu Asp Asp Cys Ala Val Leu Tyr Arg Ser Asn Ser Gln
65 70 75 80
Ser Arg Val Ile Glu Glu Ala Leu Ile Arg Cys Gln Ile Pro Tyr Arg
85 90 95
Ile Tyr Gly Gly Met Arg Phe Phe Glu Arg Gln Glu Ile Lys Asp Ala
100 105 110
Leu Ala Tyr Leu Arg Leu Ile Asn Asn Arg Gln Asp Asp Ala Ala Phe
115 120 125
Glu Arg Val Ile Asn Thr Pro Thr Arg
130 135

<210> 2053
<211> 287
<212> DNA
<213> Homo sapiens

<400> 2053
nccatggaag ccttcaatct tgtaagagaa agtgaacagc tgttttccat atgccaaatc
60
ccgctcctct gctggatcct gtgtaccagt ctgaagcaag agatgcagaa aggaaaagac
120

agttcaaaaa gttcatttat aaaagtagtt tcctgttcct agtgtgatgt atcacaaatt
2580
gtgctgaggt tattttagta tgtgtgtttc attcccgtgc ttctgttctg aagtcctgga
2640
atacagtttt cagtgtaatt aattcaactg cacttaacac taatgtccgt gttggtatag
2700
aaatgtctaa atcctatact ctagttgagg aagatcttcc ataattttat ggtattacac
2760
agggaaagct atgactgcag gatcagtcta actatactat taggtgcatg tattctcttt
2820
tcactaactt atacttgtct atctagaata caggtcttcc agtcagctgg tcatttacca
2880
ggtgtggact taagttgctg ggcttgcagt aagaattgcc agccactcat tgtgcgggtc
2940
tgcgtggagc tttaatcaga aaaagcctcc actttctgta ttatgttaac attggctcat
3000
gcatataact atctgctgct gatgtagttc tccatcttca agatttagag tgggttaacc
3060
aggtcattac atcttaattt aataacaagc attactgtag agtgattgtg tatagatctg
3120
ttagctgtca gggtgtgttt tttttaacct gttgtgtgcg tgtgggggtt aggattagta
3180
aggtgaactg ttcaggaatt ctctgcacta gctgtgcaga agagcagata actagcgctg
3240
ctctggcatt aatcccagga accactagca gtagtggggc gccgccaatc taacatgagc
3300
acaggtgctt catgacaaac attactagca tgttcaactg caccatgttc tggcactgta
3360
ttttgaatga cattaattta ttaaataaat tgtatatatt caaaaaaaaa aaaaaaaaaa
3420
aaaaaaa
3427

<210> 5002
<211> 335
<212> PRT
<213> Homo sapiens

<400> 5002
Met Ser Thr Gln Asp Glu Arg Gln Ile Asn Thr Glu Tyr Ala Val Ser
1 5 10 15
Leu Leu Glu Gln Leu Lys Leu Phe Tyr Glu Gln Gln Leu Phe Thr Asp
20 25 30
Ile Val Leu Ile Val Glu Gly Thr Glu Phe Pro Cys His Lys Met Val
35 40 45
Leu Ala Thr Cys Ser Ser Tyr Phe Arg Ala Met Phe Met Ser Gly Leu
50 55 60
Ser Glu Ser Lys Gln Thr His Val His Leu Arg Asn Val Asp Ala Ala
65 70 75 80
Thr Leu Gln Ile Ile Ile Thr Tyr Ala Tyr Thr Gly Asn Leu Ala Met
85 90 95
Asn Asp Ser Thr Val Glu Gln Leu Tyr Glu Thr Ala Cys Phe Leu Gln
100 105 110
Val Glu Asp Val Leu Gln Arg Cys Arg Glu Tyr Leu Ile Lys Lys Ile

115 120 125
Asn Ala Glu Asn Cys Val Arg Leu Leu Ser Phe Ala Asp Leu Phe Ser
130 135 140
Cys Glu Glu Leu Lys Gln Ser Ala Lys Arg Met Val Glu His Lys Phe
145 150 155 160
Thr Ala Val Tyr His Gln Asp Ala Phe Met Gln Leu Leu His Asp Leu
165 170 175
Leu Ile Asp Ile Leu Ser Ser Asp Asn Leu Asn Val Glu Lys Glu Glu
180 185 190
Thr Val Arg Glu Ala Ala Met Leu Trp Leu Glu Tyr Asn Thr Glu Ser
195 200 205
Arg Ser Gln Tyr Leu Ser Ser Val Leu Ser Gln Ile Arg Ile Asp Ala
210 215 220
Leu Ser Glu Val Thr Gln Arg Ala Trp Phe Gln Gly Leu Pro Pro Asn
225 230 235 240
Asp Lys Ser Val Val Val Gln Gly Leu Tyr Lys Ser Met Pro Lys Phe
245 250 255
Phe Lys Pro Arg Leu Gly Met Thr Lys Glu Glu Met Met Ile Phe Ile
260 265 270
Glu Ala Ser Ser Glu Asn Pro Cys Ser Leu Tyr Ser Ser Val Cys Tyr
275 280 285
Ser Pro Gln Ala Glu Lys Val Tyr Lys Leu Cys Ser Pro Pro Ala Asp
290 295 300
Leu His Lys Val Gly Thr Val Val Thr Pro Asp Asn Asp Ile Tyr Ile
305 310 315 320
Ala Gly Gly Gln Val Pro Leu Xaa Lys His Lys Asn Lys Ser Gln
325 330 335

<210> 5003
<211> 3729
<212> DNA
<213> Homo sapiens

<400> 5003
ncaggtgggc ccttgcccac cccaccctgg gaaggctggg ccaggatggg gcaggcacct
60
caccccggcc aggaacagga acgggcacca tctcggggac tgatgttttt tgaatggcgc
120
tatccaccct gccctgctcg gcctggctgt gcaggcctct tggtaccacg tctgttcgta
180
atgaccgtaa caactctatt ttcttccaca gatgactctg gggacgacga cgaggctacc
240
accccagccg acaagagcga gctgcaccac accctgaaga atctttccct gaagttagat
300
gacctcagca cgtgcaatga cctcatcgcc aagcatggcg ctgccctcca gcgctccctg
360
aatgagctgg acggcctcaa gatcccatct gagagtgggg agaagctgaa ggtggtgaat
420
gagcgggcca ccctcttccg catcacatcc aatgctatga tcaacgcctg cagggacttc
480
ttggaactag cagagataca cagtcggaaa tggcagcggg cactgcagta tgagcaggag
540
cagcgcgtgc acttggagga aaccattgag cagctggcga agcagcacaa cagcctcgag
600